UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/921,868	08/03/2001	Vikas Agarwal	JP920010088US1	7137

7590   10/31/2005

McGinn & Gibb, PLLC
2568-A Riva Road
Suite 304
Annapolis, MD 28211

EXAMINER
PATEL, ASHOKKUMAR B

ART UNIT	PAPER NUMBER
2154	

DATE MAILED: 10/31/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

Office Action Summary	Application No.	Applicant(s)
	09/921,868	AGARWAL ET AL.
	Examiner	Art Unit
	Ashok B. Patel	2154

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *03 August 2005*.

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-20* is/are pending in the application.

   4a) Of the above claim(s) *7,10-12,14,17 and 18* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-6,8,9,13,15,16,19 and 20* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

   Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

   Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   a)☐ All   b)☐ Some *   c)☐ None of:

   1.☐ Certified copies of the priority documents have been received.

   2.☐ Certified copies of the priority documents have been received in Application No. _____.

   3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

   * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
   Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413)
   Paper No(s)/Mail Date. _____.

5)☐ Notice of Informal Patent Application (PTO-152)

6)☐ Other: _____.

DETAILED ACTION

1.   Claims 1-20 are subject to examination. Claims 7, 10-12, 14, 17 and 18 have

been cancelled.

Continued Examination Under 37 CFR 1.114

2.   A request for continued examination under 37 CFR 1.114, including the fee set

forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this

application is eligible for continued examination under 37 CFR 1.114, and the fee set

forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action

has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on

08/03/2005 has been entered.

Response to Arguments

3.   Applicant's arguments filed 08/03/2005 have been fully considered but they are

not persuasive for the following reasons:

Applicant's argument:

" The virtual server in Abrams is a set of physical servers serving an application

for one customer. Whereas the virtual server provided by the Applicants' invention is

defined as a multi-tiered application which can include multiple instances of each tier

(i.e., resource classes)."

Examiner's response:

Abrams teaches at page 9, para. [0084],"Further, the virtual single server

mechanism also allows the application provider access to the appshot 220 through

conduit 360 from a single point in the on-demand system 140."

Abram also teaches at page 6, para. [0067]," In one embodiment, an appshot 220 encapsulates a multi-tier applications stack, including data 222. The present on-demand application processing method and system 140 performs this appshot encapsulation or snapshotting which saves the states of a running set of processes."

Applicant's argument:

"Conversely, the Applicant's invention provides hit-weights to control the allocation of resources in a fine-grained manner."

" Conversely, the Applicant's invention achieves that and goes one step further by optimizing the service provider's revenue in cases where multiple customer's peak occur at the same time and hence acquire more resources than available with the provider"

Conversely, the Applicants' invention allocates resources to customers based on current load and past usage history (i.e., changed number of application instances of one or more resource class components). While handling the actual incoming request the load distributor component does the following: The client of an incoming hit is identified, followed by a decision as to which subfarm application instance has enough allocated capacity allocated for the client lo be able handle the hit. This is followed by the hit being forwarded to the application."

"Moreover, the claimed invention then dynamically allocates application level resources in accordance with the requirements."

Examiner's answer:

Abrams teaches at page 15, para. [0133], "A load of tens of thousands of simultaneous hits is manageable for a single server or machine. Thus, the benefits of distributing loads becomes apparent through the scalable, on-demand capacity provided by the present system 140."

Abrams also teaches at page 2, para. [0019], [0020] and [0021]," The resources are made available upon receiving requests for a first application. Once a request is received, routing of the request is determined and the request is routed to access the first application. The application provider is then charged based on the amount of resources utilized to satisfy the request. In determining routing the method and apparatus determines if a first instance of a first application is active, and if the first instance is at a capacity. A first set of compute resources is provided to satisfy the first request and the amount charged to the first application provider is increased based on the first set of compute resources. In one embodiment, the method and apparatus activates a second instance of the first application on a second set of the available compute resources if the first instance is at capacity and the amount charged to the first application provider is increased based on the second set of compute resources. As a result, resources needed are dynamically available on demand, and freed when not needed. The application provider is only charged for services that are actually used. [0020] In one embodiment, a third set of compute resources are freed up if the compute resources are not available. A second instance of the first application is restored on a fourth set of compute resources such that the fourth set of compute resources includes at least a portion of the freed up third set of compute resources, and the amount

charged to the first application provider is increased based on the fourth set of compute resources. In freeing up resources, a first instance of a second application is snapshotted, wherein the second application is provided by a second application provider, and an amount charged to the second application provider is reduced based on the freed up third set of compute resources. [0021] The method and apparatus provides application providers with access to the network, where the network includes the distributed compute resources configured to provide the application processing and allows the application providers to distribute applications onto the network to utilize the distributed compute resources for processing of the applications. The application providers are further capable of monitoring, updating and replacing the distributed applications. The method and apparatus increases the amount of compute resources utilized in providing processing for an application as demand for the application increases. As the amount of compute resources is increased the amount charged to the application provider is increased based on the amount of compute resources utilized. As demand for the application falls, the amount of resources is reduced and the amount charged the application provider is reduced." **(provides hit-weights to control the allocation of resources in a fine-grained manner.")**

Thus Abrams does exactly what the Applicant's invention does and what claims 1, 13, 15, 16, 19 and 20 recites "providing, for each of the application-level users, respective sets of one or more application instances of each resource class component for the application on one or more machines, to service the incoming requests from respective application-level users to use the application".

Claim Rejections - 35 USC § 112

4.   The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly
> claiming the subject matter which the applicant regards as his invention.

5.   Claims 1, 13, 15, 16, 19 and 20 are rejected under 35 U.S.C. 112, second

paragraph, as being indefinite for failing to particularly point out and distinctly claim the

subject matter which applicant regards as the invention.

All these claims recite "an application comprising a plurality of resource class

components" and "respective sets of one or more application instances of each

resource class component **for the application**". The claims further recite "using hit-

weigh corresponding to a number of hits allocated for each resource instance and an

allocated weight for each resource instance".

As recited, it is unclear what the metes and bounds are of these claims. For the

purpose of this Office Action Examiner interprets these claims as being " allocating one

or more resource instances to one or more application instances as they are needed.

Claim Rejections - 35 USC § 102

6.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
> (e) the invention was described in (1) an application for patent, published under section 122(b),
> by another filed in the United States before the invention by the applicant for patent or (2) a patent
> granted on an application for patent by another filed in the United States before the invention by the
> applicant for patent, except that an international application filed under the treaty defined in section 351(a)
> shall have the effects for purposes of this subsection of an application filed in the United States only if the
> international application designated the United States and was published under Article 21(2) of such
> treaty in the English language.

7.   Claims 1-3, 5, 6, 8, 9, 13, 15, 16, 19 and 20 are rejected under 35 U.S.C. 102(e)

as being anticipated by Abrams et al. (hereinafter Abrams) (US 2002/0166117 A1)

Referring to claim 1,

Abrams teaches a method of providing access for a plurality of application-level

users to an application comprising a plurality of resource class components collectively

executing on multiple networked machines (Abstract) , the method comprising of:

'receiving an incoming flow of requests from application-level users to use an

application and components of said application (Abstract);

providing, for each of the application-level users, respective sets of one or more

application instances of each resource class component for the application on one or

more machines, to service the incoming requests from respective application-level

users to use the application (Abstract, page 2, para.[0021]);

directing each of the incoming requests to a particular application instance of an

appropriate resource class component (page 6, para.[0067]);

monitoring, for each of the application-level users, the number of request

serviced by the application instances of the resource class components of the

application (Abstract);

changing the number of application instances of one or more resource class

components in response to the monitored number of requests for each resource class

component (Abstract, page 2, para.[0021]);

maintaining a record of the current rate of requests received from respective application-level users based on the monitored number of serviced requests (page 2, para.[0021]); and

using hit-weights corresponding to a number of hits allocated for each resource instance and an allocated weight for said each resource instance, collectively and automatically allocating fractions of different resource class components to a particular application-level user in response to the changed number of application instances of one or more resource class components (page 5, para.[0062], page 6, para.[0067], page 2, para.[0019],[0020] and [0021],").

Referring to claim 2,

Abrams teaches the method as claimed in claim 1, further comprising:

directing each of the incoming requests from respective application-level users to a particular application instance of an appropriate resource class component from a respective set of one or more application instances of each resource class component, said particular application instance being identified as the least loaded of the application instances of the appropriate resource class component from that respective set. (page 6, para.[0067])

Referring to claim 3,

Abrams teaches the method as claimed in claim 1, wherein the step of providing application instances of each resource class component further comprises: initiating one or more application instance of one or more resource class on a plurality of machines to service incoming requests to use the application (page 6, para.[0067],[0068]); and

terminating one or more application instances of each resource class on a plurality of

machines to service incoming requests to use the application (page 2, para.[0021]).

Referring to claim 5,

Abrams teaches the method as claimed in claim 1, further comprising: maintaining a

record of service obligations to respective application-level users. (page 6, para.[0064],

page 14, para. [0125])

Referring to claim 6,

Abrams teaches the method as claimed in claim 5, further comprising changing, for

each of the application-level users, the number of application instances of each

resource class component in response to the monitored number of requests for each

resource class component, wherein the service obligations to respective application-

level users are at least met. (page 6, para.[0064], page 14, para. [0125], page 8,

para.[0078]).

Referring to claim 8,

Abrams teaches the method as claimed in claim 1, wherein said step of changing the

number of application instances of said one or more resource classes in (i) at least

partly based upon said recorded current rate of requests received from respective

application-level users, (page 8, para.[0074]) and (ii) at least partly based on

predetermined information that correlates changes in request rates with charges in the

corresponding number of application instances of said one or more resource classes

required to service said request rates.(page 6, para.[0068], page 8, 0078])

Referring to claim 9,

Abrams teaches the method as claimed in claim 1, wherein one or more of the

application-level users are organizations, and the requests are generated by individuals

associated with the respective organization. (page 5, para. [0059])

Referring to claim 13,

Abrams teaches the method of providing access for a plurality of application-level

users to an application comprising a plurality of resource class components collectively

executing on multiple networked machines (Abstract), the method comprising steps of:

receiving an incoming flow of requests from application-level users to use an

application and components of said application (Abstract);

providing, for each of the application-level users, respective sets of one or more

application instances of each resource class component for the application on one or

more machines, to service the incoming requests from the application-level users to use

the application (Abstract, page 2, para.[0021]);

monitoring, for each of the application-level users, the resources currently

available and resources currently consumed by the requests serviced by application

instances of the resource class components of the application (Abstract); and

maintaining a record of resources currently available to respective application-

level users; and a record of resources currently consumed by respective application-

level users; both records of said resources being maintained in respect of each of the

one or more application instances of each resource class components (page 6,

para.[0067]); and

adjusting the respective numbers of said one or more application instances of each component (Abstract, page 2, para.[0021]); and

using hit-weights corresponding to a number of hits allocated for each resource instance and an allocated weight for said each resource instance, collectively and automatically allocating fractions of different resource class components to a particular application-level user in response to a fluctuating number of application instances of one or more resource class components (page 5, para.[0062], page 6, para.[0067], page 2, para.[0019],[0020] and [0021],").

wherein said application instances of each resource class component are adjusted for each application-level user based (i) at least partly on said records of resources currently available and currently consumed by respective application-level users (page 8, para.[0074]). and (ii) at least partly on predetermined information that estimates the number of each resource class components required to service requests for said application instances of the resource class components (page 6, para.[0068], page 8, 0078]).

Referring to claim 15,

Claim 15 is a claim to a system that carries out the steps of method of claim 1. Therefore claim 15 is rejected for the reasons set forth for claim 1.

Referring to claim 16,

Claim 16 is a claim to a computer software program, recorded on a medium and capable of execution of steps of method of claim 1. Therefore claim 16 is rejected for the reasons set forth for claim 1.

Referring to claim 19,

Claim 19 is a claim to a system that carries out the steps of method of claim 13.

Therefore claim 19 is rejected for the reasons set forth for claim 13.

Referring to claim 20,

Claim 20 is a claim to a computer software program, recorded on a medium and

capable of execution of steps of method of claim 13. Therefore claim 20 is rejected for

the reasons set forth for claim 13.

Claim Rejections - 35 USC § 103

8.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

9.   Claim 4 is rejected under 35 U.S.C. 103(a) as being unpatentable over Abrams

et al. (hereinafter Abrams) (US 2002/0166117 A1) in view of Microsoft Computer

Dictionary (hereinafter Microsoft) Published in 1997.

Referring to claim 4,

Keeping in mind the teachings of Abrams, although Abrams teaches at para.[0125],

page 14, "Execution policies relate to user-level SLAs and priorities for execution.",

Abrams fails to specifically teach , wherein requests from application-level users to use

the application are stored in a queue for execution by a particular application instance of

the appropriate resource class on a first-in-first-out basis.

Microsoft teaches " a method of processing a queue, in which they were removed in the same order in which they were added – the first in is the first out.

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention was made to prioritize the execution of the requests of Abrams per Microsoft such that same user level SLAs are executed in a first-in-first-out basis.

Conclusion

Examiner's note: Examiner has cited particular columns and line numbers in the references as applied to the claims above for the convenience of the applicant. Although the specified citations are representative of the teachings of the art and are applied to the specific limitations within the individual claim, other passages and figures may apply as well. It is respectfully requested from the applicant in preparing responses, to fully consider the references in entirety as potentially teaching all or part of the claimed invention, as well as the context of the passage as taught by the prior art or disclosed by the Examiner.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ashok B. Patel whose telephone number is (571) 272-3972. The examiner can normally be reached on 8:00am-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John A. Follansbee can be reached on (571) 272-3964. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free).

Abp
***

JOHN FOU ANSBEE
SUPERVISORY PATENT EXAMINER
ART CENTER 2100